Avaliação Psicológica
ISSN 2175-3431 (versão on-line)

How Stable is the Intelligence Over Time? Evidence from Brazilian Data

Carmen Flores-Mendoza[1]
Universidade Federal de Minas Gerais – UFMG, Belo Horizonte-MG, Brasil
Daniel Marcos Andrade
Universidade Salgado de Oliveira, Belo Horizonte-MG, Brasil

ABSTRACT

There is strong scientific evidence that intelligence remains stable from childhood to adulthood; however, no study has examined the generalizability of these findings for native Brazilian samples. Here we present a study carried out in 2002 and from 2014-2017 (average timespan of 15 years) in order to verify the stability of psychometric intelligence. A sample of 120 participants (mean age Time 1=10.0 yrs, mean age Time 2=23.6 yrs) was assessed using the Raven's Progressive Matrices and verbal subtests from the WISC-III and WAIS-III. The results indicated a moderate coefficient (around .50) for intelligence differential stability independent of the measurement administered. Regarding absolute stability, real cognitive gains were observed from Time 1 to Time 2, despite controlling the effect of the regression toward the mean. Finally, sex (female) and intelligence were statistically significant predictors of rapid educational progression.
Keywords: Wechsler verbal scales; IQ; Brazil; stability.

RESUMO – Quão Estável é a Inteligência? Evidência a partir de Dados Brasileiros

Existe sólida evidência científica que a inteligência permanece estável desde a infância até a idade adulta, embora não existam estudos que verifique a generalidade desse pressuposto para amostra de nativos brasileiros. Aqui um estudo desenvolvido em 2002 e 2014/2017 (intervalo médio de 15 anos) que verifica a estabilidade da inteligência psicométrica é apresentado. Uma amostra de 120 participantes (Tempo 1=10,0 anos, Tempo 2=23,6 anos) foi avaliada com o uso das Matrizes Progressivas de Raven e as escalas verbais do WISC-III e do WAIS-III. Os resultados indicaram um coeficiente de estabilidade diferencial da inteligência moderado (aproximadamente 0,50), independentemente do tipo de medida administrada. Com relação à estabilidade absoluta, observaram-se ganhos cognitivos do Tempo 1 para o Tempo 2, apesar do controle do efeito da regressão à média. Finalmente, o sexo (feminino) e a inteligência foram os preditores estatisticamente significativos de uma rápida progressão educacional.
Palavras-chave: escala verbais Wechsler, QI, Brasil, estabilidade.

RESUMEN – ¿Cómo de Estable es la Inteligencia? Evidencias a partir de Datos Brasileños

Existen fuertes evidencias científicas de que la inteligencia permanece estable desde la niñez hasta la edad adulta; sin embargo, no hay estudios que verifiquen la generalización de este supuesto para muestra de brasileños nativos. Aquí presentamos un estudio realizado en 2002 y de 2014 a 2017 (intervalo de tiempo promedio de 15 años) que comprueba la estabilidad de la inteligencia psicométrica. Se evaluó el desempeño de 120 participantes (Tiempo 1=10,0 años y Tiempo 2=23,6 años) con las Matrices Progresivas de Raven y escalas verbales del WISC-III y del WAIS-III. Los resultados indicaron un coeficiente de la estabilidad diferencial intelectual moderada (aproximadamente .50), independiente del tipo de medida administrada. En cuanto a la estabilidad absoluta, se observaron ganancias cognitivas del Tiempo 1 al Tiempo 2, a pesar de controlar el efecto de la regresión hacia la media. Por último, el sexo (femenino) y la inteligencia fueron predictores estadísticamente significativos de una rápida progresión educativa.
Palabras clave: Escalas verbales Wechsler; IQ; Brasil; estabilidad.

Intelligence is the most successfully investigated construct in psychological science. Moreover, considering its ubiquity in a broad array of life outcomes the construct is one of the most important contributions of psychology to modern society (Lynn, & Vanhanen, 2012; Warne, 2020). The robust empirical findings coming from the neuroimaging and behavioral genetic research permit intelligence to be considered mainly, but not exclusively, a biological variable (Bruton, 2021; Jensen, 1998; Haier, 2016; Hilger et al., 2022) and, conjecturally, to be considered as product of evolutionary process (Bruner & Colom, 2022). However, intelligence research must meet several methodological requirements (e.g., valid universal measures, representative samples,

[1] Endereço para correspondência: Universidade Federal de Minas Gerais. Faculdade de Filosofia e Ciências Humanas – Gabinete 4042. Avenida Antônio Carlos, 6627, 31270-901. Belo Horizonte, MG. *E-mail:* carmenflor@uol.com.br
Paper derived from the doctoral thesis by Daniel de Andrade, supervised by the first author (Carmen Flores-Mendoza), defended in 2018 in the postgraduate program in Developmental Psychology at the Federal University of Minas Gerais.

many-labs studies, longitudinal studies, etc.) to confirm these assumptions. Defined as the general ability for reasoning, abstract thinking, learning from the experience, understanding complex ideas, and solving problems (Gottfredson, 1997), intelligence is related to achievements throughout the life-course. In this regard, whatever the measure used (verbal or non-verbal), as long as it requires high cognitive process, empirical research indicates that individuals who score highly on cognitive tests tend to achieve better school performance/educational attainment, both individually and at the group level (Firkowska-Mankiewicz, 2011; Hegelund et al., 2018; Lynn & Vanhanen, 2012; Zisman & Ganzach, 2022; Warne, 2020), and they tend to be less involved in crime, prolonged training periods, legal issues, alcoholism, car accidents (Lynn & Vanhanen, 2012), in addition to presenting a lower risk of all-cause mortality (Cukic et al., 2017; Warne, 2020).

Nevertheless, note that associations do not warrant causal interpretations of the relationship between intelligence and life outcomes because is not clear which are the mediating/moderator variables that explains or affect the strength and direction of that relationship. In this regard, SES and personality are the intermediate variables most cited by the academy. However, contradictory results are found. For instance, von Stumm and Plomin (2015) analyzed the developmental relationship between SES (measured by parental education and occupation) and intelligence (using parent-administered tests, ratings of abilities, and a mixture of web-based, telephone-based, and parent-administered tests) of 14,853 twin children, cognitively assessed nine times between 2 and 16 years. The authors found that the difference in IQ points at age 2 between high and low SES families tripled at age 16, favoring children from high SES. In this case, to disentangle the SES-intelligence dyad seems to be no possible. This type of reciprocal effect has been named "Matthew effects" (those who begin with advantage/disadvantage accumulate more advantage/disadvantage over time). However, in another study, O'Connell and Marks (2020) analyzed longitudinal data from 6216 children. The cognitive ability was measured through numerical, verbal and overall reasoning ability. The results indicated that the cognitive ability at age 13 was the stronger predictor of academic attainment at age 16 than SES measures (composed by occupational class, household income, and parental education). In this case, SES did not account for the variability in academic attainment.

To further complicate the issue, there is no conclusive evidence that SES is a variable exclusively representative of the environment. Perhaps, SES is a result of the interaction gene x environment (Trzaskowski et al., 2014). In this regard, Marks and O'Connell (2021) analyzed the cognitive performance of children (through the Peabody Picture Vocabulary test, digit span memory, reading comprehension, reading recognition and math) of the 1979 National Longitudinal Study of Youth mothers (NLSY79). The results indicated small increases in the SES-cognitive performance correlations over time; however, the effects of SES dropped when the analysis considered the mother's ability and the prior ability of the children. Thus, the "Mattew effects" could be explained by confounding variables such as mother's abilities. Moreover, the SES effect on intelligence seems change as children grow older (moderate influence in childhood and no influence in adulthood) (Gottschling et al., 2019).

Regarding personality, Borghans et al. (2016) analyzed four datasets (including the NLSY79 database), arriving to the conclusion that dimensions of personality have greater predictive power than IQ for achievement test scores, and grades (school attainment), which, in turn, are better predictors of important life outcomes (e.g. wage, welfare, physical health). As the natural corollary of these results, personality would be more important than IQ for predicting life outcomes. A replication of this study was conducted by Zisman and Ganzach (2022) with two databases used by Borghans et al. (2016), and four additional databases. The results indicated an average R2 .232 of intelligence and .053 of personality for educational attainment. The same trend for occupational success was found. Thus, Zisman and Ganzach's results were the opposite of those Borghans et al. (2016). As there is no conclusive answer on the nature and the influence of SES or personality on the individual's life, we assume that the relationship between intelligence and life outcomes could be genuine; however, the moderator/mediator factors in this relationship still are unknown.

Identifying the relationship between intelligence and social outcomes was possible thanks to the effort of psychologists by creating accurate psychological tests over a little more than a century of research. Despite controversial debates about the psychometric structure of intelligence and between-test comparability (Bünger et al., 2021), most intelligence tests used in modern research and applied psychology show strong evidence for reliability and validity following international guidelines such as the International Test Commission (2017) or local guidelines such as the Brazilian Psychological Test Assessment System-SATEPSI (https://satepsi.cfp.org.br/).

In this scenario, it is important to know how stable the individual differences in cognitive ability are from childhood to adulthood as measured by traditional intelligence tests. This knowledge impacts the cognitive assessment activity of psychologist on diverse settings, especially on clinical and educational context. In this regard, IQ cross-sectional and long-term studies has provided solid findings on intelligence stability (Deary, 2014, 2015; Hunt, 2010; Ramsden et al., 2011), both at the level of differential stability, which refers to the changes in participants within a population, and at the level of absolute stability, which refers to gains or loss across time.

For instance, at differential level, Sameroff et al. (1993), using the Wechsler Scales, found a correlation of .72 between cognitive assessment at four years old and retesting at 13 years old. Rönnlund et al. (2015), using different cognitive measures, found a latent regression of β=.95 between intelligence scores at age 18 and 50. Schalke et al. (2013) found a mean correlation of .80 between cognitive measures administered to 344 adults (56% females) assessed in 1968 (age mean=12) and 2008 (age mean=52). In identical and fraternal twins, Plomin et al. (1994) found a phenotypic stability of .93 in participants reared apart and reared together. Lyons et al. (2009) found a correlation of .74 between cognitive assessments of 1237 male twins at 20 and 55 years old.

Additionally, some studies found higher correlations at earlier ages, while others identified higher correlations at later ages. Aimed to clarify these discrepancies, the Seattle Longitudinal Study (SLS) was designed and conducted by then-doctoral student Warner Schaie in 1956 (Shaie & Willis, 2010). After several waves of data collection, retesting previous samples, and adding new random samples, Shaie and collaborators found strong correlations between cognitive measures administered in individuals at different ages (for youth a *r* of .88; middle-age a *r* of .90, and advanced age a *r* of .94). However, the interval between testings would be the key to understand the variety of stability coefficients. For example, for near assessments Gustafsson and Undhein (1992) found a high correlation (.92) for 12-15 yrs old; Schneider et al. (2014) found *r* of .95 for 17-23 yrs old; Yu et al. (2018) found *r* of .91 for infancy-preschool, and Breit et al. (2021) found differential coefficients of specific abilities between .72 and .84 for an interval of 6-month. On the contrary, reduced correlations were observed between farther ages. For instance, Sameroff et al. (1993) found a *r* of .72 for 4-13 yrs old; Schneider et al (2014) found a *r* of .40 for 4-17 yrs and *r* of .46 for 4-23 yrs; and Yu et al. (2018) found a r of .57 for infancy-adolescence.

Therefore, stability coefficients may decrease as the timespan between assessment waves increase. In this regard, the estimate of Conley (1984), based on several longitudinal consistency studies, could be correct. He estimated an annual stability of intelligence of .99, which decreases to .95, .90, .82, .74, and .67 for intervals of 5, 10, 20, 30 and 40 years, assuming perfect instrument reliability.

Currently, the most cited project regarding IQ stability over an extended interval is the Scottish Mental Survey led by the British psychologist Ian Deary. It was a follow-up study to the Scottish Mental Survey from 1932 and 1947 when almost all Scottish 11-year-olds were tested for their thinking abilities. The investigation is considered the most extended follow-up study of intelligence. Several reports were published from the Scottish Mental Survey (Deary, 2014; Deary & Brett, 2015; Deary et al., 2013). The general result indicated that about half of the differences in intelligence at older age ($.70^2$) could be traced back to childhood.

Most of studies are conducted at IQ level, and not many at g level (also known as general level) is found. One of these studies was conducted by Larsen et al. (2008). They analyzed approximately 4000 members from the Vietnam Experience Study, which were first cognitively assessed in 1967-1971 (mean age=19.9 years), then 18 years later (mean age=38.3 years). During the initial testing, five cognitive tests were administered, while the second testing administered 14 cognitive tests. A Principal Axis Factor analysis was conducted separately for Time 1 and Time 2 in order to extract a general factor (*g*) that represented a general cognitive ability. The differential stability coefficients were .85 for g; .79 for arithmetic; and .82 for verbal ability. Regarding absolute stability, an increase of 9.4 IQ points in verbal ability, but no increase in arithmetic were observed. Also, Ronnlund et al. (2015) examined a small sample of males (*n*=262) assessed at age 18 (three tests used as indicators of *g*), at age 50 (four measures), and then at five-year intervals up to age 65 (four measures). Results from structural equation modeling indicated standardized regression coefficients of .95 between young-adult *g* and the *g* factor at ages 50-55, and a slight decrease at following ages (.94 at age 60, and .86 at age 65). The mean limitation in stability research at the latent level is the few tests used for representing the *g* factor (Larsen et al. used five and 14 tests, and Ronnlund et al. used three and four tests for time 1 and time 2). In this regard, Jensen (1998) indicated at least nine cognitive tests would be necessary for obtaining a reasonable g. Moreover, the battery of tests should be in the same quantity and content in time 1 and time 2, which is extremely difficult to administer in a long follow-up study.

However, as observed, all studies show positive correlations between cognitive measures during all ages of the life cycle, corroborating the assumption that intelligence (at general or specific abilities) is a stable psychological construct, including early ages, such as 1 and 2 years old (Breit, et al., 2020, 2021; Girault, et al., 2018).

Regarding absolute stability of intellectual ability over time (i.e., whether children with initially low and high IQ scores remain in the same zone when they grow up), Schneider et al. (2014) found that 50% of children at age 4 and 60% at age 7 classified as high, average, and low IQ were reclassified in the same subgroups later on in life. In general, these results showed that most people tend to maintain their cognitive functioning over time; however, some IQ changes may happen with advancing years, mainly when the first cognitive assessment was conducted early. Breit et al. (2021), in their study of high ability students, found an increase of 8.8 IQ points (from 116.7 to 125.5) for general ability and, for specific abilities, the gains ranged from 5.6 IQ points (reasoning) to 10.3 IQ points (figural ability). Considering individual-level change, 30.7% showed an increase in IQ; in terms of specific abilities 5.3% increased their creativity (the smallest increase), and 23.7% increased their processing speed (the largest increase). These results indicated that in high cognitive performance, robust stability is found. On the other hand, the meta-analysis conducted by Whitaker (2008) on people with low IQ also indicated strong differential stability coefficients (.77 and .78 for Verbal IQ and Performance IQ, respectively), but in terms of absolute stability, 14% of the Full-Scale IQ changed by 10 points or more. Therefore, the IQ cross-sectional and longitudinal studies evidenced that individuals tend to hold the same intelligence position relative to an age cohort; however, some significant changes can happen depending on the intellectual level of people.

Conversely, in Latin America, the study conducted by Mansukoski et al. (2019) found a mean correlation coefficient of .348 between infancy/adolescence (5 to 17 yrs) and elderly age (64 yrs or above) in a sample of 42 high socioeconomic status Guatemalans born 1941-1953. The authors concluded that the results were evidence of IQ instability. Violence, high crime rates, substantial political and social instability in Guatemala were considered events that could have influenced the variability of the test scores over time. However, the authors acknowledge that, beyond the small sample, the main limitation of their study was the set of different tests used in the follow-up study. In Brazil, to our knowledge, there are no studies that examined the stability of intelligence differences from childhood to early adult age using the same family of cognitive measures. The present study aims to verify differential and absolute intellectual ability stability at two-time points (average of 15 years).

Method

Participants

The cognitive performance of one hundred and twenty young adults, aged between 19 and 28 years (M=23.6, SD=2.3), 56.7% males, was analyzed in the present study. The participants were assessed during 2014-2017 (79% between 2015-2016). The same sample was evaluated in 2002 when they were between the ages of 6 and 15 years (M=10.0, SD=2.3). Thus, an average timespan between assessments of 15 years. This sample was part of the study, "Longitudinal project of Intelligence and Personality", which was approved in 2002 by the Ethical Committee CAAE: 17793814.9.0000.5149. Here forward, the 2002 sample will be referred to as Time 1, and the sample from 2014-2017 will be referred to as Time 2. The characteristics of the studied sample are in Table 1.

Table 1
Sociodemographic Characteristics of the Sample Studied at Two Time Points

Time points	Education level	*f*	Age		SES*
			M	*SD*	*f*
Time 1	1° grade	23	7.2	.45	A1-A2=6 B1=23 B2=25 C1-C2=39 D-E=3
	2° grade	9	8.1	.33	
	3° grade	27	9.3	.47	
	4° grade	19	10.1	.45	
	5° grade	12	11.4	.53	
	6° grade	8	12.6	.77	
	7° grade	12	13.3	.47	
	8° grade	10	14.2	.40	
Time 2	HS (incomplete)	1	23.0	-	A1-A2=20 B1=25 B2=47 C1-C2=27 D-E=1
	HS (complete)	12	23.1	2.9	
	UG (incomplete)	74	22.5	1.5	
	UG (complete)	22	25.7	1.9	
	GRA (incomplete)	10	26.0	1.9	
	GRA (complete)	1	27	-	

Note. *information of socioeconomic status (SES) in Time 1 was available only for 96 participants. HS=High School; UG=Undergraduation; GRA=Graduation

Instruments

Wechsler verbal scales (Wechsler, 2002 and 2004). The verbal scale of the Wechsler Adult Intelligence Scale (WAIS-III) and of the Wechsler Intelligence Scale for Children (WISC-III) were administered at the young adult age and child age, respectively. The verbal scale was composed of Vocabulary, Information, Similarities, Comprehension, Digit Span, and Arithmetic (for adults, Letter-Number Sequencing is added). The WISC-III was published in Brazil with norms for the Brazilian context in 2002, the same year of data collection for Time 1. The WAIS-III was published in 2004, and its norms were valid for the Brazilian context until the data collection of Time 2.

Matrices Progressives of Raven (Angelini et al., 1999; Raven et al., 1998). The Standard and Coloured Matrices Progressive of Raven were administered (SPM and CPM, respectively). The CPM was administered in 2002.

Social Economic Status (SES). The researchers used The *Criterio Brasil* (www.abep.org) to assess and categorize the socioeconomic status of the participants' families. The *Criterio Brasil* was considered the most reliable classifier for describing Brazilian socioeconomic stratification compared to other classification algorithms (Kamakura & Mazzon, 2016). There was information of *Criterio Brasil* for 96 participants of Time 1 obtained from the assessment of 2004. In this time there were six categories (A, B1, B2, C1, C2, D-E), while *Criterio Brasil* for Time 2 had seven (A1, A2, B1, B2, C, D, E). For this reason, we combined categories D and E (Time 1), and C1 and C2 (Time 2). The final SES scale for Time 1 and Time 2 was D-E=1, C1-C2=2, B2=3, B1=4, and A1 and A2=5. The A category was representative of the highest socioeconomic status and D-E representative of the lowest socioeconomic status. Note that participants in Time 2 still resided with their parents or were financially dependent on them.

Procedure

The dataset analyzed here is based on the Longitudinal Study of Intelligence and Personality conducted by the Laboratory of Individual Differences at the Federal University of Minas Gerais (authorization number 17793814.9.0000.5149). The study began in 2002 recruiting almost 650 schoolchildren and they were assessed every two years until 2010, for both intellectual and personality variables. In 2014, the participants' recruitment was performed by the second author accessing telephone numbers and physical addresses that were filed in the primary school where the first assessment was conducted. The search was extended to secondary school, where most students attended after finishing primary school. However, it was not possible to contact most of the students assessed in 2002 due to the lack of updated information regarding address and/or phone number. Thus, the strategy was adapted to asking the contacted participants to provide updated information about their peers who attended the same primary school in 2002. This updated information permitted contacting other participants by phone, and all information regarding the project was provided (objectives, test performance time, access to their 2002 results). The participation was voluntary. No financial compensation was provided. The testing activities were conducted by the authors and three psychologists from Monday to Saturday according to the participants' availability. Each testing session took an average of 4 hours, including a rest period of 20 minutes, where a small snack was provided by the project. In Time 1, participants were assessed in quiet rooms at the school where the study was conducted, following the parents' signed research participation agreement. In Time 2, all participants were assessed individually in quiet rooms at the university, after the research participation agreement forms were signed. For ages 7-9 years, the CPM was administered in small groups (3-5 participants). The SPM was administered collectively in Time 1, and individually in Time 2. The WISC-III and the WAIS-II verbal scales were administered individually in Time 1 and Time 2. In both periods (Time 1 and Time 2), a team of psychology students and psychologists were trained by the first author in proper administration of the psychological tests.

In Time 1, four participants did not respond to the SPM test, and three participants did not respond to the WISC-III Verbal Scale. In Time 2, all participants responded to all cognitive measures. Thus, from 120 assessed participants, we analyzed 116 SPM protocols and 117 WISC-III verbal scale protocols (see Figure 1).

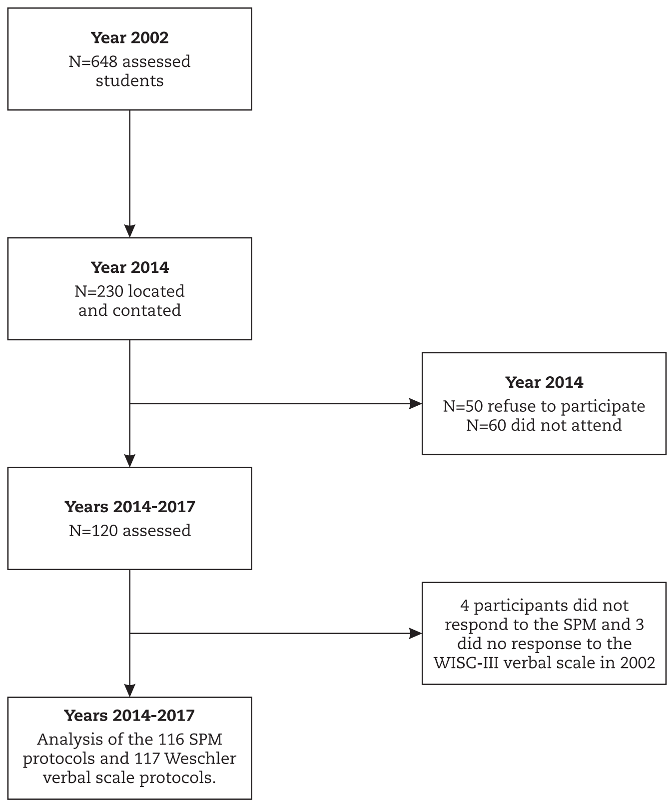
Data analysis

Our study refers to observations conducted at two-time points; thus, the present study cannot be considered a longitudinal study (Singer, & Willett, 2003). For observations at two-time points, change score and correlation/regressions techniques are appropriated (Garcia, & Marder, 2017). The correlation is an estimation of differential stability. The repeated-measures t-test estimates absolute stability (Breit, et al., 2022).

To better understand the differences between differential and absolute stability, here is an example. If there were 10 subjects that presented IQ 80, 82, 90, 82, 110, 111, 105, 105, 119, and 125 in 1990 (Time 1), and 20 years after (Time 2), the same subjects showed IQ 88, 70, 80, 80, 100, 115, 100, 90, 120, and 125, respectively, the correlation in this example would be .914 (r) or .908 (r_{ho}); a high positive correlation. This result refers to differential stability. However, note in the same example that some participants had cognitive losses or gains of up to 15 IQ points (or one standard deviation). This information refers to absolute stability. Therefore, given the former, the latter does not necessarily follow. Differential and absolute stability are two independent pieces of information, and both add information to the subject of cognitive stability. At this point, it is worth noting the words of Ian Deary (the researcher who leads the Scottish Mental Survey): “The correlations should not be taken at face value” (2014, p. 242). The reason is that various factors contribute to the residual non-stable variance such as the participant’s testing practice, error of the measurement, restriction of the cognitive range of the studied sample, changes in participant's health over time, and others. Thus, an almost perfect differential or absolute stability coefficient is not expected.

Pearson's correlations and repeated-measures test-*t* were conducted on our normally distributed variables [Kolmogorov-Smirnov test D(116)=.069, p=.200 for SPM Time 1, and D(116)=.055, p=.200 for SPM Time 2. In cases of verbal scales D(117)=.081, p=.067 for Time 1, and D(117)=.066, p=.200 for Time 2]. We used the SPSS software version 22 (IBM, 2013) throughout this project.

Figure 1
Recruitment Flowchart



Standardized scores were used throughout the process. In the SPM test case, most of the participants from Time 1 (*N*=84) responded to the CPM instead of the SPM test due to their young age. In these cases, the CPM raw scores were converted into SPM raw scores using Table 27 of the Raven's Manual (Raven et al., 1998). All the raw scores of the SPM test (Time 1 and Time 2) were converted to percentiles according to the norms elaborated for the state of Minas Gerais (Flores-Mendoza et al., 2014), the same region of origin of the participants in this study. Then the percentile ranks were converted to an IQ metric. Regarding WISC-III/WAIS-III verbal scales, the IQ scores were estimated according to the Brazilian manuals (Wechsler, 2002, 2004).

Regarding the absolute stability, we conducted a repeated measure test-t on IQ scores of the SPM and WISC-III verbal scale. Additionally, we estimated changes of IQ categories for both cognitive measures. For this purpose, five IQ categories were defined (1=Very Low <IQ 84; 2=Low IQ 85 – IQ 99; 3=Middle IQ 100 – IQ 114; 4=Superior IQ 115 – IQ 129, and 5=Very Superior >IQ 130) for Time 1 and Time 2. Differences between IQ categories were classified as gainers if the adult IQ category was superior to the child IQ category, maintainers if the adult IQ category was the same in childhood, or losers if the adult IQ category was inferior to the child IQ category.

Finally, regression analysis was conducted to verify the contribution of potential predictors (child and adult cognitive performance, child and adult SES, and sex) on fast-forward educational progression, a variable represented through a special elaborated scale. Note that fast-forward educational progression was preferred over educational attainment due to our sample enrollment (64%) or due to those that had already finished university (25%). Our scale matched age and the expected school level to the participant's chronological age. The idea was to present a scoring bonus to participants with more schooling at a younger age. For example, a 26-year-old participant with a maximum education level equivalent to incomplete high school would receive 1 point. However, if a 20-year-old participant reached post-graduation, the participant would receive 7 points. The lower the age and the higher the education, the more points the participant would receive on the achievement scale. The predictors were chosen based on the hypothesis that SES could add incremental validity to the prediction for fast-forward educational progression beyond cognitive performance. It is known that intelligent people would accelerate their education more than less smart people because smart people do not need extended time to reach knowledge. However, this assertion would be correct if all participants had the same socioeconomic level for enrolling in courses, training, college or universities; which was not true in our Brazilian sample. For this reason, the SES variable was introduced into our regression model. Regarding sex, it was hypothesized that if females enroll more than males in educational settings, females probably differ from males in fast-forward educational attainment. For this reason, sex was introduced into our regression model.

The association between the fast-forward educational progression scale and IQ categories at Time 1 was .289 (p=.002). Low IQ performers had a mean of 3.0, while high IQ performers had 5.35 points in educational progression.

Results

Differential stability

The correlation SPM/Verbal Scale for Time 1 was .386 (p=.000; or corrected correlation after removing the effect of random measurement error of .435), and .641 (p=.000; or corrected correlation of .694) for Time 2. Therefore, the correlation between cognitive measures was lower in childhood than in adulthood. The correlation coefficient between SPM Time 1 and SPM Time 2 was .496 or a corrected correlation of .548. The correlation between the verbal scales of WISC-III and WAIS-III was .407 (or .490 corrected correlation). Therefore, the differential stability coefficient for non-verbal cognitive measures was slightly higher than for verbal ones. The results are shown in Table 2.

Table 2
Correlation Between Cognitive Performance at Time 1 (T1) and Time 2 (T2)

	Verbal IQ T1	Verbal IQ T2	SPM IQ T1	SPM IQ T2
Verbal IQ T1	1			
Verbal IQ T2	.407**	1		
SPM IQ T1	.386**	.370**	1	
SPM IQ T2	.284**	.641**	.496**	1

Note. **Correlation statistically significant at .01 level

Absolute stability

The mean SPM IQ was 109 (SD=10.4) and 100 (SD=13.4) for Time 1 and Time 2, respectively. A repeated-measures t-test found this difference to be statistically significant [t(115)=6.922, p=.000] between Time 1 and Time 2. The differences in IQ points in each category is shown in Figure 2. While category 1 gained 11.8 IQ points, category 4 lost 10.9 IQ points. In terms of percentages, 68.5% of gainers were concentrated in categories 1 and 2, and 75.5% of losers were concentrated in category 4. However, the regression toward the mean could explain this result (see Discussion).

Regarding the Wechsler verbal scale, the mean verbal IQ was 115 (SD=11.9) and 112 (SD=8.3) for Time 1 and Time 2, respectively. A repeated-measures t-test found this difference to be statistically significant [t(116)=3.299, p=.001]. The differences in IQ points in each category is shown in Figure 2. While category 1 and 2 gained an average of 21 IQ points, category 5 lost an average of 11 IQ points. In terms of percentages, 80% of gainers were category 1 and 2, and 93% of losers were concentrated in categories 4 and 5. However, the regression toward the mean could be a partial explanation for this result (see Discussion).

Fast-forward education progression

Multiple linear regression analysis was used to develop a model for predicting fast-forward education progression. Basic descriptive and regression coefficients are shown in Table 3. Only sex and childhood cognitive performance had a statistically significant (p<.01) zero-order correlation with educational attainment, and both had statistically significant (p<.05) partial effects in the full model. The predictor model was able to account for 13% of the variance in fast educational attainment, F(5, 90)=2.600, p<.030, R^2=.127; 95% CI [.17, .52].

The results in Table 3 indicated that to be female and the cognitive performance at the childhood period predicted early adulthood fast-forward education progression in our sample. IQ adulthood and SES were not significant predictor variables.

Figure 2
Mean Differences in IQ Points for the SPM Test (Left Side) and for the Wechsler Verbal Scales (Right Side) Without Controlling for the Regression toward the Mean

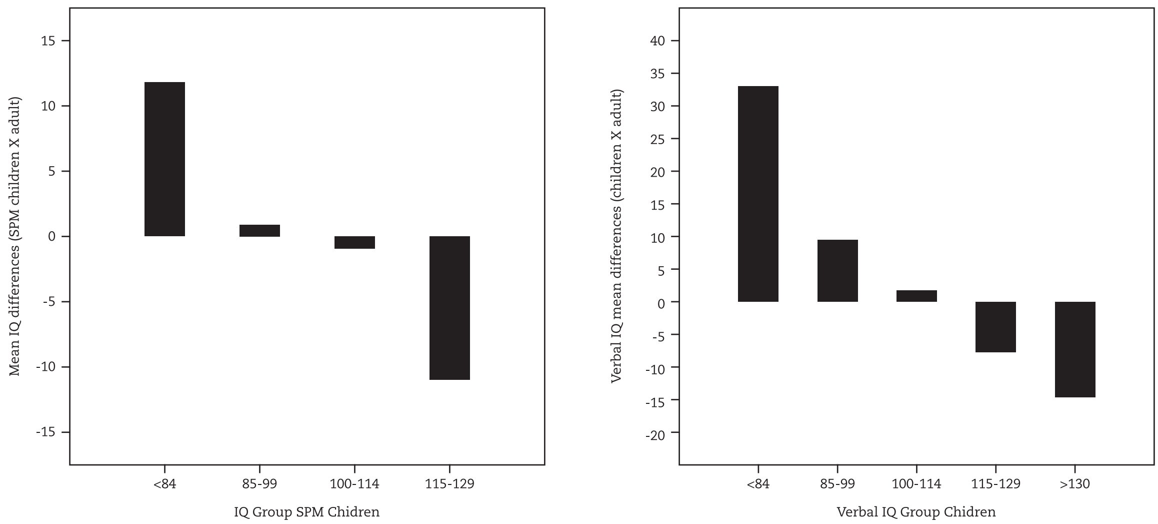


Table 3
Fast-forward Education Progression in Adulthood Related to Verbal Intelligence Performance and Socioeconomic Status (N=96)

Variable	Zero-Order *r*					β	*sr*	*b*
	IQ T1	IQ T2	SES T1	SES T2	FEP			
SES T1					.107	.241	.159	.167
SES T2				.449**	.044	-267	.177	.193
IQ T2			.212*	.210*	.117	.005	.029	.030
IQ T1		.379**	.178*	.135	.224*	.026	.218	.226
Sex	.019	.020	.087	.104	.194*	.664	.229	.233
							Intercept=1.740	
Mean	114.5	111.2	2.89	3.34	4.9			
SD	12.4	8.0	1.01	1.06	1.4		R^2=.127*	

Discussion

We intended to verify the relative stability of cognitive performance test scores in two-time points (average time interval=15 years) using the two most reliable and well-known intelligence tests (the SPM test and the Wechsler Verbal Scales). Four clear-cut results were obtained.

Differential stability. Our results indicated once again that intelligence is, at least, a moderate stable psychological factor. However, the stability coefficients found in the present study (around .50) are not as high as those found in some other studies (Deary, & Brett, 2015; Deary et al., 2013). The long time interval between the two assessment moments could explain the results obtained. Previous studies (Schneider et al., 2014) indicated lower correlations between the farthest closest ages (more distal) than closest ages (more proximal). In addition, due to biological maturation reasons, genetic effects on IQ differ from infancy (30%) to adulthood (60-70%) (Jensen, 1998). In other words, when we compare childhood cognitive performance with adulthood cognitive performance (farthest ages), we are comparing an initial performance that responds strongly under the environmental influence with a later performance that reacts under genetic influence.

Different stability coefficients. There is no consensus in the literature about differences in coefficients of stability for measures that are less culturally influenced, as is the case with the SPM test, and for measures that receive strong cultural influences, such as with the Wechsler verbal scales. While some studies found high differential stability coefficients for measures that are highly influenced by education and culture (Deary & Brett, 2015; Ramsden, et al., 2011; Salthouse, 2014; Schwartzman et al., 1987), others found high coefficients for measures for culturally less influenced (Larson et al, 2008). Corroborating this last finding, the meta-analysis conducted by Scharfen et al. (2018) indicated that the size of the retest effects, also considered an indicative of stability, differed between tests that used different kinds of contents. Fluid intelligence (kind of intelligence less culturally influenced) showed higher stability than crystallized intelligence (kind of intelligence highly influenced by education and culture). Moreover, while the effect declined with a longer test-retest interval between administrations, ability level or age were not significant moderators of the stability. Our study indicated that the SPM test showed a little higher stability (corrected r=.548) than the verbal scales (corrected r=.490) on a timespan average of 15 years. Still, this difference between non-verbal and verbal stability coefficients was non statistically significant (p=.277). Note that even using cognitive measures families (CPM/SPM; WISC-III/WAIS-III), which can increase the chance of error to the estimations, the correlations obtained between the measures were still positive and moderate. Additionally, we found a lower correlation between the SPM test and the Wechsler verbal scale at Time 1 (.386 or .435 after correction for range attenuation) compared to Time 2 (.641 or .694 after correction for restriction range). This result can be in accordance with the dedifferentiation hypothesis, which states that an increase in the correlations across abilities is driven by increasing age (Blum & Holling, 2017; Hartung et al.; Hülür et al., 2015). At the same time, it reveals that there might not be a homogeneous cognitive structure throughout the ages.

The stability coefficients obtained in our study are higher than those found in the Guatemalan study (Mansukoski et al., 2019), and it would be attractive to hypothesize the reasons for such a difference. However, the last study was characterized by several limitations such as being composed of a small sample (N=42), high SES, and the use of different instruments for Time 1 and Time 2, which challenge the validity of its results.

Absolute stability and the phenomenon of regression toward the mean. We found that cognitive gainers concentrated (68% in the case of the SPM test to 80% for WISC-III verbal scale) in categories of low cognitive performance, and conversely, cognitive losers (75%-SPM to 93%-WISC-III verbal scale) concentrated in categories of high cognitive performance. Before concluding that children with poor cognitive performance achieve a higher cognitive performance in adulthood or that children with high cognitive performance lose cognitive performance in adulthood is necessary to understand the well-known statistical phenomenon known as regression toward the mean, which receives insufficient attention in social and behavioral psychological research (Yu & Chen, 2015).

The phenomenon refers to the tendency of scores to average out, and it happens for two reasons. The first reason is when a correlation between two measures is less than imperfect (in our study the r between Time 1 and Time 2 was .496, and .407 for the SPM test, and the WISC-III verbal scale respectively without attenuation for range restriction). That means that some higher performers did well on the first try (Time 1) and did worse on their second try (Time 2) and vice versa. This is partially due to an error in measurement. The second reason is due to the use of a nonrandom sample from a population. In our case, a convenient sample was studied. It was composed of participants who accepted to be assessed during their young adult age. The mean verbal IQ was 115 and 112 for Time 1 and Time 2, respectively. These averages were slightly higher than the school population, where the sample was recruited (which was 110). If our sample were a random sample with a mean IQ of 110, there would be no regression towards the population's mean.

The following formula is used to estimate the regression's percent to the mean:

$$Prm = 100\ (1-r)$$

According to this formula, we obtained 50%, and 60% of the regression to the mean for the SPM test, and the WISC-III verbal scale, respectively. In other words, a high percentage of changes was due to the statistical phenomenon.

At this point, another question arises: is it possible to calculate how many pseudo-points of IQ were due to regression to the mean? The answer is, 'Yes'. For a simple calculation (see details in Trochim, 2022), we need three pieces of information: (a) the mean IQ of the population (in our case a verbal mean IQ of 110); (b) the verbal mean IQ of the sample studied (88 for low performers, and 127 for high performers), and (c) the percentage of regression (60% in case of the WISC-III verbal scale). Considering low performers and the Wechsler Verbal IQ, there was a 13-point pseudo-effect (60% of the way from IQ 88 to IQ 110). For high performers, there was a 10-point pseudo-effect (60% from IQ 127 to IQ 110). Thus, if we subtract the 13 pseudo-points from the 21 verbal IQ points that low performers supposedly gained, then the actual gain of this group was 8 verbal IQ points. On the other hand, subtracting a 10-point pseudo-effect from the 11 verbal IQ points that high performers supposedly lost would indicate that this group had only a decrease of 1 verbal IQ point. To estimate the percentage of changes and pseudo-IQ points due to regression toward the mean in intelligence stability research has implications for diagnosis and classifications in clinical settings. For instance, the estimated initial gain of 21 verbal IQ points for lower performers is equivalent to 1.4 standard deviations. If this estimate was accurate, clinical psychologists can arrive at the wrong conclusion that any cognitive diagnosis or classification in childhood would be unreliable. However, instead of an unexpected gain of 21 IQ points, the real estimated gain was 8 IQ points (after controlling the regression toward the mean) a gain somewhat expected, considering the average effect of education (Ritchie & Tucker-Drob, 2018) or the effect of cognitive training (Protzko, 2016; Ree & Carretta, 2022) on IQ. Nevertheless, we recognize that the gain cognitive observed exclusively in initially low-performing individuals (also described by Schroeders et al., 2016) deserves further investigation.

Prediction. Our dependent variable, "fast-forward education progression," was not related to educational attainment, variable with small variability (almost 90% of participants were enrolled or had entered university). It was associated with the early achievement of higher levels of education considering chronological age. For this, a particular scale was developed (see Analysis). The results indicated that socioeconomic status was not a statistically significant predictor, while sex (female) and intelligence were statistically significant predictors of educational attainment. In this regard, international surveys (Organisation for Economic Co-operation and Development, 2019) indicate that females tend to keep pursuing education more than men (e.g., women make up 55% of upper secondary graduates and 54% of post-secondary non-tertiary graduates). Therefore, it is more likely to observe that women, instead of men, reach a higher educational level at a younger age. On the other hand, it is well known that smarter individuals tend to achieve higher academic degrees in less time than their less intelligent peers (McClarty, 2015). Therefore, on this point, our results were unsurprising; however, corroborate the predictive strength of intelligence.

Conclusion

Intelligence is one of the most frequently investigated phenotypes in behavioral science. To the best of our knowledge, there is no study indicating how stable intelligence is when measured in the Brazilian context at two distant points of time. Our study replicated previous results found in intelligence research. Whatever measure (SPM or WISC-III verbal scale) were considered, intelligence appears as a moderately stable psychological construct, and it predicts social outcomes as educational progression. Thus, our study provides evidence that adulthood cognitive function and scholastic mobility can be predicted from childhood IQ.

Additionally, it was possible to observe that positive changes (IQ gains) are more accentuated at the intelligence's lower tail. This result could be a clue for the investigation of another socio-behavioral phenomenon, called the Flynn Effect, which is observed in large groups. Low IQ countries are gaining IQ points in each subsequent generation, whereas high IQ countries are losing IQ points (Rindermann et al., 2017). However, we recognize that our sample was small and non-representative of the general population, and most of the observed IQ classification variation could be explained by the statistical phenomenon of regression toward the mean. Despite this, there was room for some real cognitive gain in the lower tail. Future investigation with new cohorts will provide more precise answers.

Acknowledgments

There are no mentions.

Funding

This work was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) N° 472181/2013-0.

Authors’ contributions

We declare that all authors participated in the preparation of the manuscript. Specifically, author Carmen Flores-Mendoza is the general coordinator of the study, who designed the study, re-analyzed the data, trained the data collection team, and wrote the article. Daniel Andrade coordinated data collection and participated in the final writing of the work.

Availability of data and materials

All data and syntax generated and analyzed during this research will be treated with complete confidentiality due to the Ethics Committee for Research in Human Beings requirements. However, the dataset and syntax that support the conclusions of this article are available upon reasonable request to the principal author of the study.

Competing interests

The authors declare that there are no conflicts of interest.

References

Angelini, A. L., Alves, I .C. B., Custódio, E. M., Duarte, W. F., & Duarte, J. L. M. (1999). *Matrizes Progressivas Colorias de Raven – Escala Especial.* São Paulo: Centro Editor de Testes e Pesquisas em Psicologia.

Blum, D., & Holling, H. (2017). Spearman's law of diminishing returns. A meta-analysis. *Intelligence, 65*, 60-66. http://dx.doi.org/10.1016/j.intell.2017.07.004.

Breit, M., Brunner, M., & Preckel, F. (2020). General intelligence and specific cognitive abilities in adolescence: Tests of developmental differentiation, ability differentiation, and their interaction in two large samples. *Developmental Psychology, 56*(2), 364-384. https://doi.org/10.1037/dev0000876.

Breit, M., Sherrer, V., Preckel, F. (2021). Temporal stability of specific ability scores and intelligences profiles in high ability students. *Intelligence, 86*, 101538. https://doi.org/10.1016/j.intell.2021.101538.

Bruton,O.J. (2021). Is there a "g-neuron"? Establishing a systematic link between general intelligence (g) and the von Economo neuron. *Intelligence, 86*. https://doi.org/10.1016/j.intell.2021.101540.

Bünger, A., Grieder, S., Schweizer, F., & Grob, A. (2021). The comparability of intelligence test results: Group- and individual-level comparisons of seven intelligence tests. *Journal of School Psychology, 88*, 101-117. https://doi.org/10.1016/j.jsp.2021.09.002.

Conley, J. (1984). The hierarchy of consistency: A review and model of longitudinal findings on adult individual differences in intelligence, personality and self-opinion. *Personality and Individual Differences, 5*(1), 11-25. https://doi.org/10.1016/0191-8869(84)90133-8.

Čukić, I., Brett, C.E., Calvin, C.M., Batty, G.D., & Deary, I.J. (2017). Childhood IQ and survival to 79: Follow-up of 94% of the Scottish Mental Survey 1947. *Intelligence, 63*, 45-50. https://doi.org/10.1016/j.intell.2017.05.002.

Deary, I. J., & Brett, C. E. (2015). Predicting and retrodicting intelligence between childhood and old age in the 6-Day Sample of the Scottish Mental Survey 1947. *Intelligence, 50*, 1-9. https://doi:10.1016/j.intell.2015.02.002.

Deary, I. J., Pattie, A., & Starr, J. M. (2013). The stability of intelligence from age 11 to age 90 years: the Lothian birth cohort of 1921. *Psychological Science, 24*(12), 2361-2368. https://doi:10.1177/0956797613486487.

Deary, I.J. (2014). A estabilidade da inteligência desde a infância até a velhice. *Current Directions in Psychological Science, 23*(4), 239-245. https://doi.org/10.1177/0963721414536905.

Firkowska-Mankiewicz, A. (2011). Adult Careers: Does Childhood IQ Predict Later Life Outcome? *Journal of Policy and Practice in Intellectual Disabilities*, 8, 1, 1-9.

Flores-Mendoza, C., Widaman, K. F., Bacelar, T. D., & Lelé, A. J. (2014). Propriedades psicométricas do Raven Geral no contexto de Minas Gerais. *Arquivos Brasileiros de Psicologia, 66*(2), 1-16.

Garcia T.P., Marder K. (2017). Statistical Approaches to Longitudinal Data Analysis in Neurodegenerative Diseases: Huntington's Disease as a Model. *Current Neurology and Neuroscience Reports, 17*(14). https://doi.org/10.1007/s11910-017-0723-4.

Girault, J.B., Langworthy, B.W., Goldman, B. D., Stephens, R.L., Cornea, E., Reznick, J.S., Fine, J., & Gilmore, J. H. (2018). The predictive value of developmental assessments at 1 and 2 for intelligence quotients at 6. *Intelligence, 68*, 58-65. https://doi.org/10.1016/j.intell.2018.03.003.

Gottfredson, L.S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence, 24*, 13-23. https://doi:10.1016/S0160-2896(97)90011-8.

Gottschling, J. Hahn, E., Beam, C.R., Spinath, F.M., Carroll, S., & Turkheimer, E. (2019). Socioeconomic status amplifies genetic effects in middle childhood in a large German twin sample. *Intelligence, 72*, 20-27. https://doi.org/10.1016/j.intell.2018.11.006.

Haier, R.J. (2016). *The Neuroscience of Intelligence.* Cambridge University Press.

Hartung, J., Doebler, P., Schroeders, U., & Wilhelm, O. (2018). Dedifferentiation and differentiation of intelligence in adults across age and years of education. *Intelligence, 69*, 37-49. https://doi.org/10.1016/j.intell.2018.04.003.

Hegelund, E. R., Flensborg-Madsen, T., Dammeyer, J., Mortensen, E.L. (2018). Low IQ as a predictor of unsuccessful educational and occupational achievement: A register-based study of 1,098,742 men in Denmark 1968-2016. *Intelligence, 71*, 46-53. https://doi.org/10.1016/j.intell.2018.10.002.

Hülür, G., Ram, N., Willis, S. L., Schaie, K. W., & Gerstof, D. (2015). Cognitive dedifferentiation with increasing age and proximity of death: Within-person evidence from the Seattle Longitudinal Study. *Psychology and Aging, 30*, 311-323. https://doi.org/10.1037/a0039260.

IBM Corp. Released 2013. *IBM SPSS Statistics for Windows, Version 22.0.* Armonk, NY: IBM Corp.

International Test Commission. (2017). *The ITC Guidelines for Translating and Adapting Tests* (Second edition). [www.InTestCom.org].

Jensen, A. R. (1998). *The g factor.* Praeger.

Kamakura, W. & Mazzon, J. A. (2016). Critérios de Estratificação e Comparação de Classificadores Socioeconômicos no Brasil. *Revista de Administração de Empresas [online], 56*(1), 55-70. https://doi.org/10.1590/S0034-759020160106.

Larsen, L., Hartmann, P., & Nyborg, H. (2008). The stability of general intelligence from early adulthood to middle-age. *Intelligence, 36*(1), 29-34. http://doi: 10.1016/j.intell.2007.01.001.

Lynn, R., & Vanhanen, T. (2012). *Intelligence: A unifying construct for the social sciences.* Ulster Institute for Social Research.

Lyons M. J., York T. P., Franz C. E., Grant M. D., Eaves L. J., Jacobson K. C., … Kremen W. S. (2009). Genes determine stability and the environment determines change in cognitive ability during 35 years of adulthood. *Psychological Science, 20*(9), 1146-1152. https://doi.org/10.1111/j.1467-9280.2009.02425.x.

Mansukoski, L., Hogervorst, E., Fúrlan, L., Galvez-Sobral, J. A., Brooke-Wavell, K., & Bogin, B. (2019). Instability in longitudinal childhood IQ scores of Guatemalan high SES individuals born between 1941-1953. *PloS One, 14*(4), e0215828. https://doi.org/10.1371/journal.pone.0215828.

McClarty, K. L. (2015). Life in the Fast Lane: Effects of Early Grade Acceleration on High School and College Outcomes. *Gifted Child Quarterly, 59*, 3-13. https://doi.org/10.1177/0016986214559595.

O'Connell, M. & Marks, G.N. (2020). Are the effects of intelligence on student achievement and well-being largely functions of family income and social class? Evidence from a longitudinal study of Irish adolescents. *Intelligence, 84*, 101511. https://doi.org/10.1016/j.intell.2020.101511.

Organization for Economic Co-operation and Development. (2019). *Education at a Glance 2019: OECD Indicators.* OECD Publishing, Paris, https://doi.org/10.1787/f8d7880d-en.

Plomin, R., Pedersen, N. L., Lichtenstein, P., & McClearn, G. E. (1994). Variability and stability in cognitive abilities are largely genetic later in life. *Behavior Genetics, 24*(3), 207-215. https://doi.org/10.1007/BF01067188.

Ramsden, S., Richardson, F. M., Josse, G., Thomas, M. S., Ellis, C., Shakeshaft, C., Seghier, M. L., & Price, C. J. (2011). Verbal and non-verbal intelligence changes in the teenage brain. *Nature, 479* (7371), 113-116. https://doi.org/10.1038/nature10514.

Raven, J.C.; Court, J.H & Raven, J. (1998). *General overview* (1999 edition). Oxford Psychologists Press.

Ree, M.J. & Carretta, T.R. (2022). Thirty years of research on general and specific abilities: Still not much more than *g*. *Intelligence, 91*, 101617. https://doi.org/10.1016/j.intell.2021.101617.

Rindermann, H., Becker, D., & Coyle, T.R. (2017). Survey of expert opinion on intelligence: The FLynn effect and the future of intelligence. *Personality and Individual Differences, 106*, 242-247. https://doi.org/10.1016/j.paid.2016.10.061.

Ritche, S.J. & Tucker-Drob, E.M. (2018). How much does education improve intelligence? A meta-analysis. *Psychological Science, 29*, 1358-1369. https://doi.org/10.1177/0956797618774253.
Rönnlund, M., Sundström, A., & Nilsson, L.G. (2015). Interindividual differences in general cognitive ability from age 18 to age 65 years are extremely stable and strongly associated with working memory capacity. *Intelligence, 53*, 59-64. https://doi.org/10.1016/j.intell.2015.08.011.
Salthouse, T.A. (2004). Localizing age-related individual differences in a hierarchical structure. *Intelligence, 32*, 541-561. https://doi.org/10.1016/j.intell.2004.07.003.
Sameroff, A. J., Seifer, R., Baldwin, A., & Baldwin, C. (1993). Stability of intelligence from preschool to adolescence: The influence of social and family risk factors. *Child Development, 64*(1), 80-97. https://doi.org/10.2307/1131438.
Schaie, K. W., & Willis, S. L. (2010). The Seattle Longitudinal Study of Adult Cognitive Development. *ISSBD Bulletin, 57*(1), 24-29. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3607395/.
Scharfen, J., Peters, J.M., & Holling, H. (2018). Retest effects in cognitive ability tests: A meta-analysis. *Intelligence, 67,* 44-66. https://doi.org/10.1016/j.intell.2018.01.003.
Schalke, D., Brunner, M., Geiser, C., Preckel, F.. Keller, U., Spengler, M., & Martin, R. (2013). Stability and change in intelligence from age 12 to age 52: Results from the Luxembourg MAGRIP study. *Developmental Psychology, 49*(8), 1529-1543. https://doi: 10.1037/a0030623.
Schneider, W., Niklas, F., & Schmiedeler, S. (2014). Intellectual development from early childhood to early adulthood: The impact of early IQ differences on stability and change over time. Learning and Individual *Differences, 32*, 156-162. https://doi: 10.1016/j.lindif.2014.02.001.
Schroeders, U., Schipolowski, S., Zettler, I., Golle, J., & Wilhelm, O. (2016). Do the smart get smarter? Development of fluid and crystallized intelligence in 3rd grade. *Intelligence, 59*, 84-95. https://doi.org/10.1016/j.intell.2016.08.003.
Schwartzman, A. E., Gold, D., Andres, D., Arbuckle, T. Y., & Chaikelson, J. (1987). Stability of intelligence: A 40-year follow-up. Canadian *Journal of Psychology/Revue Canadienne de Psychologie, 41*(2), 244-256. https://doi.org/10.1037/h0084155.
Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* Oxford University Press, USA.
Trochim, W.M.K. (2022, January, 10). *Regression to the Mean.* https://conjointly.com/kb/regression-to-the-mean/.
Trzaskowski, M., Harlaar, Arden, R., Krapohl, E., Rimfeld, K., McMillan, A., Dale, P.S., & Plomin, R. (2014). Genetic influence on Family socioeconomic status and children's intelligence, *Intelligence, 42*, 83-88. http://dx.doi.org/10.1016/j.intell.2013.11.002.
von Stumm, S. & Plomin, R. (2015). Socioeconomic status and the growth of intelligence from infancy through adolescence. *Intelligence, 48*, 30-36. https://doi.org/10.1016/j.intell.2014.10.002.
Trochim, W.M.K. (2022, January, 10). *Regression to the Mean.* https://conjointly.com/kb/regression-to-the-mean/.
Warne, R.T. (2020). *In the know: Debunking 35 myths about human intelligence.* Cambridge University Press.
Wechsler, D. (2002). *WISC: escala de inteligência para crianças: manual (3a.ed.). Adaptação e padronização para uma amostra brasileira realizada por Vera Lúcia Marques de Figueiredo.* Casa do Psicólogo.
Wechsler, D. (2004). *WAIS-III: Manual para administração e avaliação. Adaptação e padronização para uma amostra brasileira realizada por Elizabeth do Nascimento.* Casa do Psicólogo.
Whitaker, S. (2008). The Stability of IQ in People with Low Intellectual Ability: An Analysis of the Literature. *Intellectual and Developmental Disabilities, 46*(2), 120-128. http://eprints.hud.ac.uk/id/eprint/4283/.
Yu, H., Mccoach, D. B., Gottfried, A. W., & Gottfried, A. E (2018). Stability from infancy through adolescence: An autoregressive latent variable model. *Intelligence, 69*, 8-15. https://doi.org/10.1016/j.intell.2018.03.011.
Yu, R., & Chen, L. (2015). The need to control for regression to the mean in social psychology studies. *Frontiers in Psychology, 5*, 1574. https://doi.org/10.3389/fpsyg.2014.01574.

recebido em maio de 2022
aprovado em setembro de 2022

Sobre os autores

Carmen Flores-Mendoza is a professor and researcher at the Federal University of Minas Gerais. She finished her doctoral program at the University of São Paulo and pursued her postdoc at the University of California, Davis.

Daniel Marcos Andrade is a professor at Universidade Salgado de Oliveira, Campus Belo Horizonte and finished his doctoral program at the Universidade Federal de Minas Gerais – UFMG.

Como citar este artigo

Flores-Mendoza, F., & Andrade, D. M. (2022). How stable is the intelligence over time? Evidence from Brazilian data. *Avaliação Psicológica, 21*(4), 371-382. http://dx.doi.org/10.15689/ap.2022.2104.24041.01